

Exponentially modified Gaussian and the Lognormal for common physics fitting.

Milind Diwan

3/25/2020

Some notes and derivations.

Gaussian-modified-exponential

A normally distributed random number with an addition of an exponential random number is called an exponential-Gaussian or Gaussian-exponential.

The characteristic function is

$$\varphi_{EG}(s) = \frac{e^{isq_0} e^{-\frac{1}{2}\sigma^2 s^2}}{(1 - is/c)}$$

The PDF that corresponds to this is

$$P_{EG}(x) = \frac{c}{2} e^{\frac{c^2 \sigma^2}{2}} e^{-c(x-q_0)} \text{Erfc}\left[\frac{1}{\sqrt{2}}\left(c\sigma - \frac{x-q_0}{\sigma}\right)\right]$$

Recall that c is the constant for exponential PDF: $\theta(x)c e^{-cx}$, σ is the std. dev. of the Gaussian q_0 is the mean of the Gaussian.

$\text{Erfc}[x]$ is the complement of the error function.

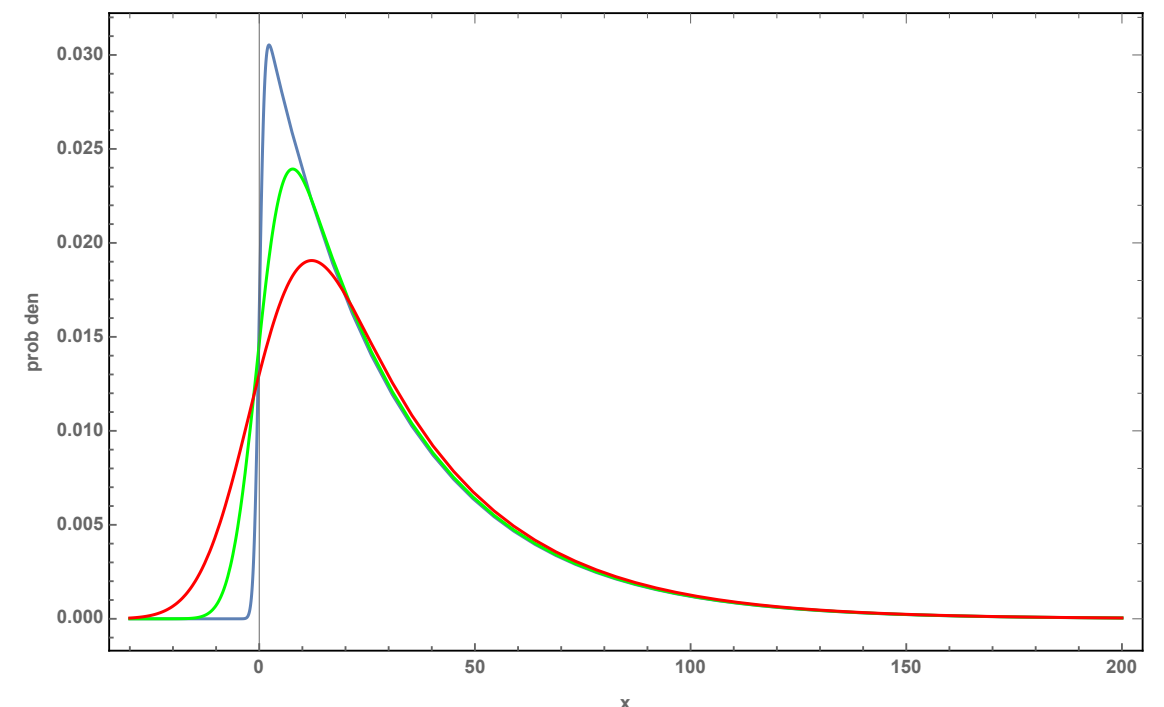
$$\text{Erfc}[x] = 1 - \text{Erf}[x] = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$$

When $c\sigma$ is small the Erfc acts like a step function.

Some care is needed in calculation in case of negative or very large arguments.

The Mean of the PDF is $(q_0 + 1/c)$

The Variance is $(\sigma^2 + 1/c^2)$



Calculate the exp-normal

To calculate the PDF we could invert the characteristic function.

$$P_{EMG}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{isq_0} e^{-\frac{1}{2}\sigma^2 s^2}}{(1 - is/c)} e^{-isx} ds$$

But in this case the explicit convolution integral seems to be easier. We use the fact that the exponential is defined above 0 to our advantage. Set $q_0 = 0$ for ease.

$$P_{EMG}(x) = \int_{-\infty}^{\infty} \theta(z) \cdot c \cdot e^{-cz} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-x)^2}{2\sigma^2}} dz = \frac{c}{\sqrt{2\pi\sigma^2}} \int_0^{\infty} e^{-cz} e^{-\frac{(z-x)^2}{2\sigma^2}} dz$$

We now complete the square in the power of the exponential.

$$P_{EMG}(x) = \frac{c}{\sqrt{2\pi\sigma^2}} \int_0^{\infty} dz \cdot e^{-\frac{1}{2\sigma^2}[(z-(x-c\sigma^2))^2 + 2xc\sigma^2 - c^2\sigma^4]}$$

$$P_{EMG}(x) = \frac{c}{\sqrt{2\pi\sigma^2}} e^{-xc} e^{c^2\sigma^2/2} \int_{-(y-c\sigma^2)/\sqrt{2}\sigma}^{\infty} \sqrt{2}\sigma du e^{-u^2}$$

$$P_{EMG}(x) = \frac{c}{2} e^{-xc} e^{c^2\sigma^2/2} \text{Erfc}\left[\frac{1}{\sqrt{2}}\left(c\sigma - \frac{x}{\sigma}\right)\right]$$

using the definition of $\text{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$

Calculate the exp-normal CDF

It is useful to calculate the CDF also. This allows ease of Monte Carlo simulation, and calculation of confidence levels.

$$F_{EMG}(y) = \frac{c}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^y dx \int_0^{\infty} e^{-cz} e^{-\frac{(z-x)^2}{2\sigma^2}} dz$$

To do this integral we have to carefully examine the domain of integration and the order of integration. First we switch the order of integration and change variables $t = (x - z)$

$$F_{EMG}(y) = \frac{c}{\sqrt{2\pi\sigma^2}} \int_0^y \int_{-\infty}^{\infty} e^{-cz} e^{-\frac{(z-x)^2}{2\sigma^2}} dx dz = \frac{c}{\sqrt{2\pi\sigma^2}} \int_0^{y-z} \int_{-\infty}^{\infty} e^{-cz} e^{-\frac{t^2}{2\sigma^2}} dt dz$$

Notice that the integral over t has an upper limit to $y - z$, and the integral over z is over $(0, y - t)$. and so we can switch the integration and change limits.

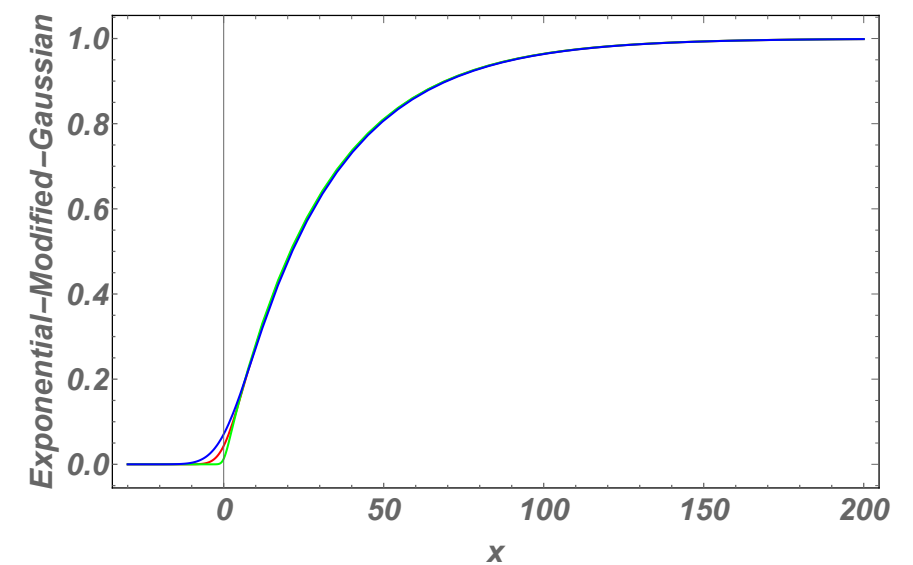
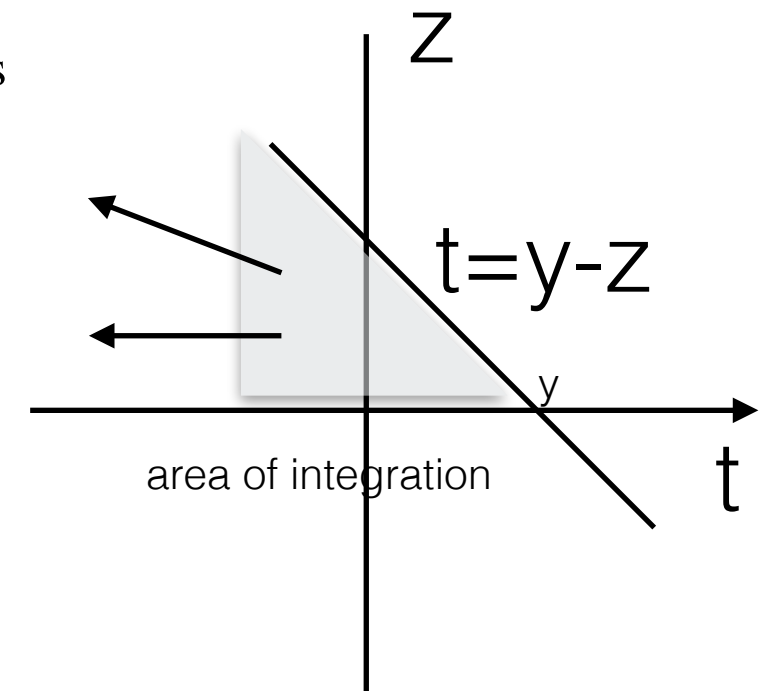
$$F_{EMG}(y) = \frac{c}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^y \int_{-\infty}^{y-t} e^{-cz} e^{-\frac{t^2}{2\sigma^2}} dz dt$$

and now we are integrating only over a exponential in the z direction.

$$F_{EMG}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^y e^{-\frac{t^2}{2\sigma^2}} (1 - e^{-c(y-t)}) dt$$

For the second term we again to the trick of completing the square for the exponent

$$F_{EMG}(y) = \frac{1}{2} (1 + \text{Erf}(\frac{y}{\sqrt{2}\sigma}) - \frac{1}{2} e^{-c(y - \frac{c\sigma^2}{2})} (1 + \text{Erf}(\frac{1}{\sqrt{2}\sigma}(y - c\sigma^2)))$$



lognormal distribution

The normal distribution arises as we take the sum of repeated measurements or random numbers drawn from some distribution.

The sum tends to a normal distribution. $X = \sum_{i=1}^n x_i$

In this sum, the random numbers have support over $(-\infty, +\infty)$.

We can ask how is the exponential of X distributed. $Y = e^X$

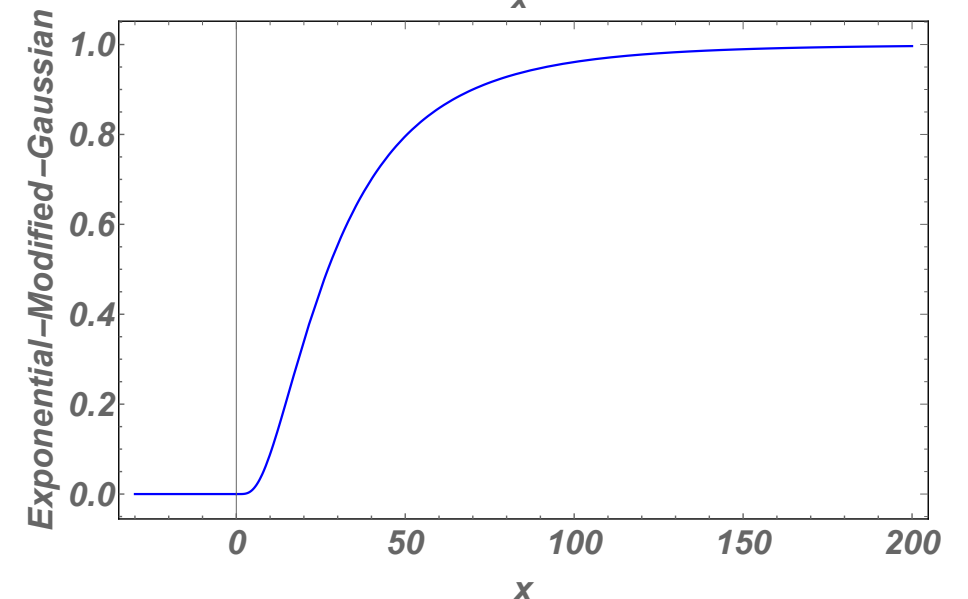
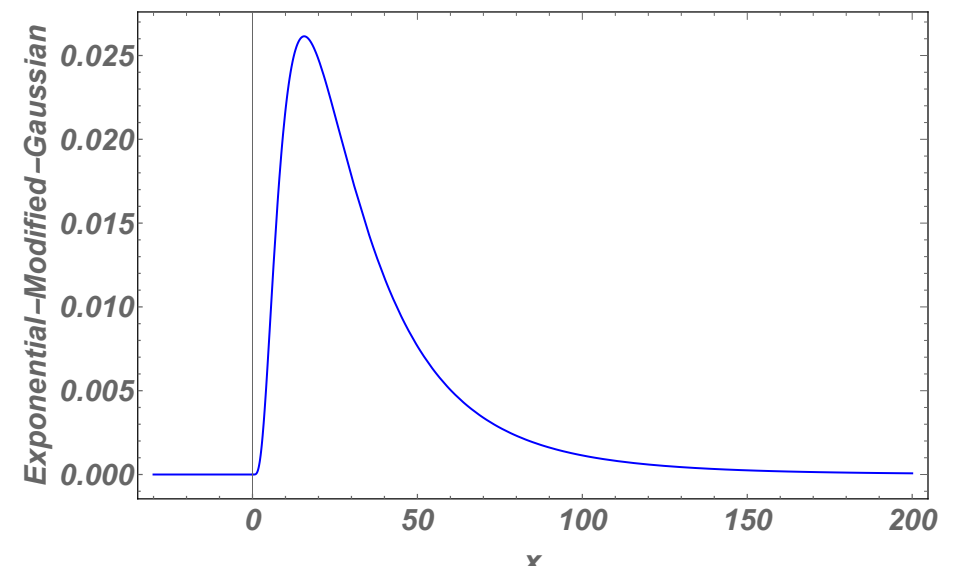
$Y = \prod_{i=1}^n y_i$ where $y_i = e^{x_i}$. However now $y_i > 0$

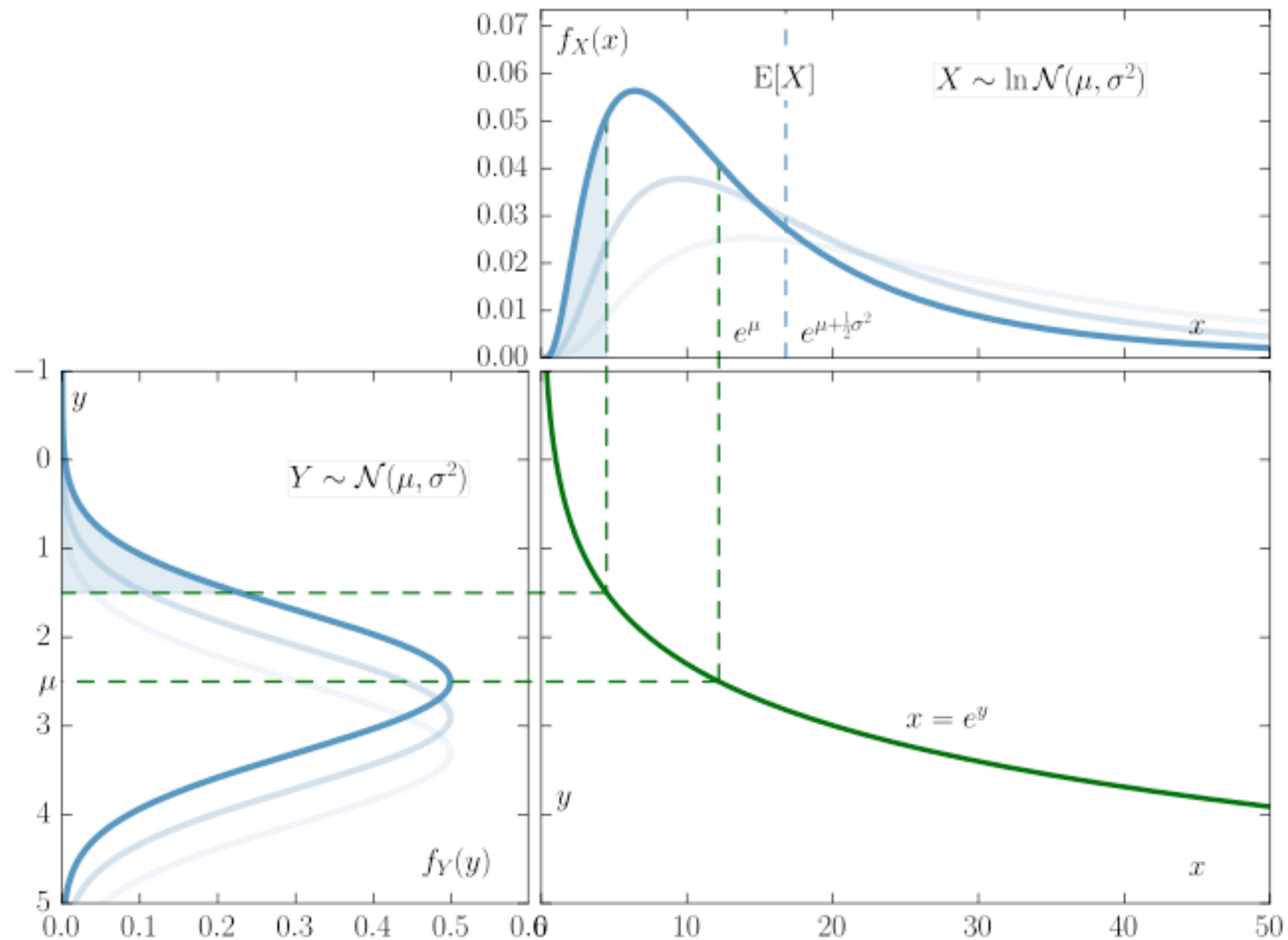
$$P_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \dots \text{ use } P_Y(y)dy = P_X(x)dx$$

$$P_Y(y) = \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot y} e^{-\frac{(\text{Log}(y)-\mu)^2}{2\sigma^2}} \quad \dots \quad y > 0$$

The cumulative distribution function is easy to show

$$F_Y(y) = \frac{1}{2} \left[1 + \text{erf} \left(\frac{\text{Log}(y) - \mu}{\sqrt{2}\sigma} \right) \right]$$





Y is normally distributed and so $X = e^Y$ will be lognormal distributed. An important distinction is, of course, that X ranges from 0 to ∞ .

lognormal moments and characteristic function

The moments can be calculated by integrating

$$\langle x^n \rangle = \int_0^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot x} e^{-\frac{(\text{Log}[x]-\mu)^2}{2\sigma^2}} x^n dx$$

$$\langle x^n \rangle = \int_0^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot x} e^{-\frac{(\text{Log}[x]-\mu)^2}{2\sigma^2} + n\text{Log}[x]} dx \quad \dots \text{complete the square}$$

$$\langle x^n \rangle = \int_0^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot x} e^{-\frac{(\text{Log}[x]-(\mu+n\sigma^2))^2}{2\sigma^2} + (\mu n + n^2\sigma^2/2)} dx = e^{n\mu + n^2\sigma^2/2}$$

For the characteristic function there appears to be no closed formula

$$\phi_X(s) = \int_0^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot x} e^{-\frac{(\text{Log}[x]-\mu)^2}{2\sigma^2}} e^{i \cdot x \cdot s} dx \quad \dots \text{if } s \text{ has a negative imaginary part then this diverges}$$

But we know the moments and so they could be used to construct a CF in an infinite series.

$$\phi_X(s) = \sum_{n=0}^{\infty} \frac{(is)^n}{n!} e^{n\mu + n^2\sigma^2/2} \dots \text{It can be proved that this diverges also.}$$

Here is everything on one page.

	Normal	Expnormal	Lognormal
PDF	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{c}{2} e^{\frac{c^2\sigma^2}{2}} e^{-c(x-\mu)} \text{Erfc}\left[\frac{1}{\sqrt{2}}\left(c\sigma - \frac{x-\mu}{\sigma}\right)\right]$	$\frac{1}{\sqrt{2\pi}\cdot\sigma\cdot(x)} e^{-\frac{(\text{Log}(x)-\mu)^2}{2\sigma^2}}$
Support	$(-\infty, \infty)$	$(-\infty, \infty)$	$(x) \in (0, \infty)$
CDF	$\frac{1}{2}(1 + \text{Erf}(\frac{x-\mu}{\sqrt{2}\sigma}))$	$\frac{1}{2}(1 + \text{Erf}(\frac{x-\mu}{\sqrt{2}\sigma})) - \frac{1}{2} e^{-c((x-\mu)-\frac{c\sigma^2}{2})} (1 + \text{Erf}(\frac{1}{\sqrt{2}\sigma}((x-\mu)-c\sigma^2)))$	$\frac{1}{2}[1 + \text{erf}(\frac{\text{Log}(x)-\mu}{\sqrt{2}\sigma})]$
CF	$e^{i\mu s - \sigma^2 s^2 / 2}$	$\frac{e^{is\mu} e^{-\frac{1}{2}\sigma^2 s^2}}{(1 - is / c)}$	$\sum_{n=0}^{\infty} \frac{(is)^n}{n!} e^{n\mu + n^2\sigma^2 / 2} \dots \text{divergent}$
Mean	μ	$\mu + 1 / c$	$e^{\mu + \sigma^2 / 2}$
Var	σ^2	$\sigma^2 + 1 / c^2$	$(e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$
Central Moments	$\langle (x - \mu)^p \rangle = \sigma^p \prod_{k=1}^{p/2} (2k - 1) \text{ for } p \text{ even}$ $0 \text{ for } p \text{ odd}$	$\langle (x - \mu)^p \rangle =$ $1 / c \dots p = 1$ $2 / c^2 + \sigma^2 \dots p = 2$ $6 / c^3 + 3\sigma^2 / c \dots p = 3$	$\langle x^n \rangle = e^{n\mu + n^2\sigma^2 / 2}$

The lognormal arithmetic moments are provided. The lognormal could be shifted by a parameter (M). This will shift the mean by M.

exercise to make lognormal mimic an exp-normal

Start with an Exp-normal.

$$\mu=0, c=1/30, \sigma=6$$

$$1/c=30,$$

$$2/c^2 + \sigma^2 = 1836$$

$$6/c^3 + 3\sigma^2/c = 165240$$

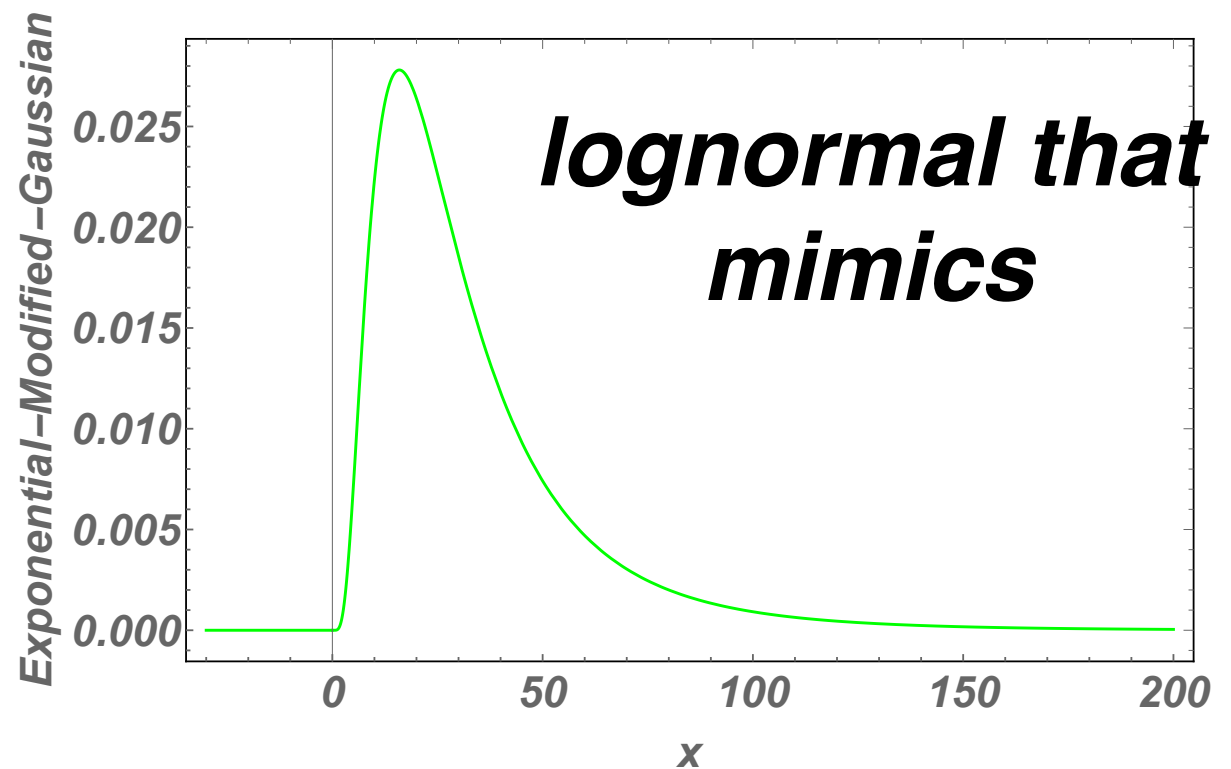
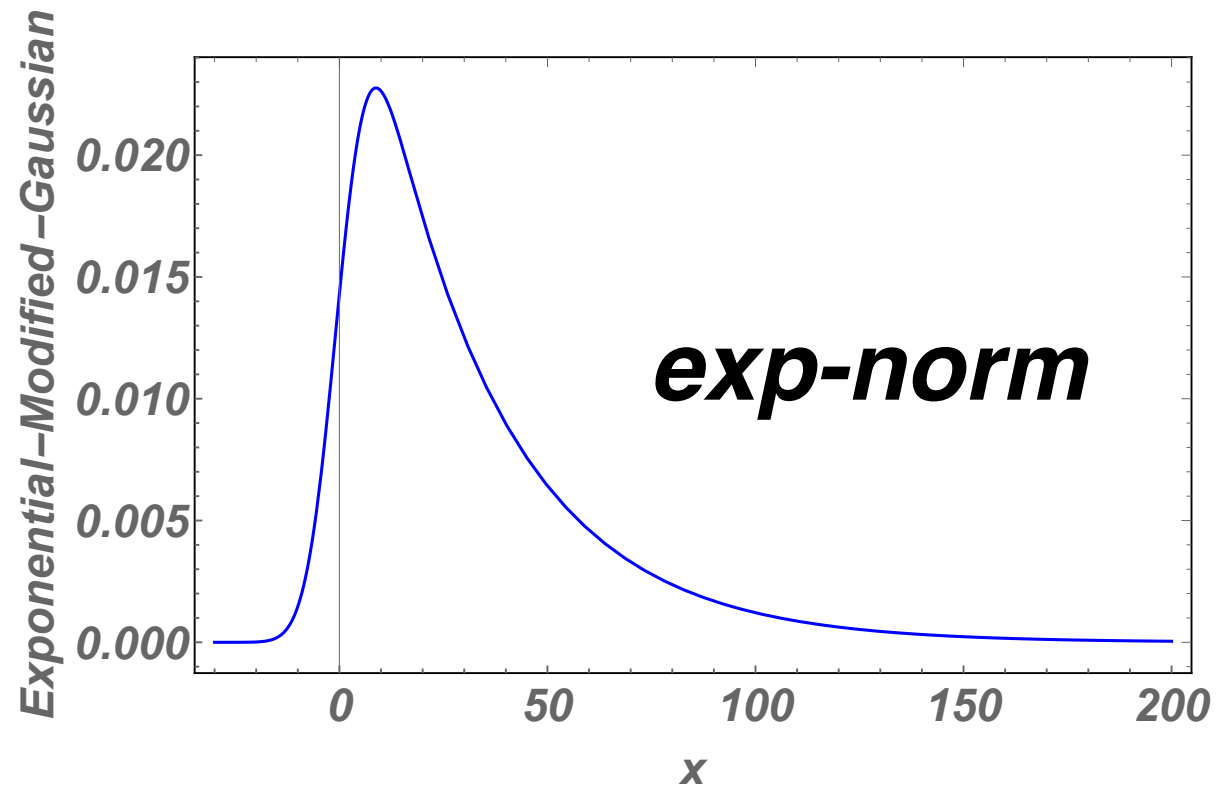
match the 2nd and 3rd moment

$$2\mu + 2\sigma^2 = \text{Log}(1836) = A$$

$$3\mu + 9\sigma^2/2 = \text{Log}(165240) = B$$

$$\Rightarrow \mu = \frac{1}{6}(9A - 4B) = 3.26291$$

$$\Rightarrow \sigma^2 = \frac{1}{3}(-3A + 2B) = 0.494758$$

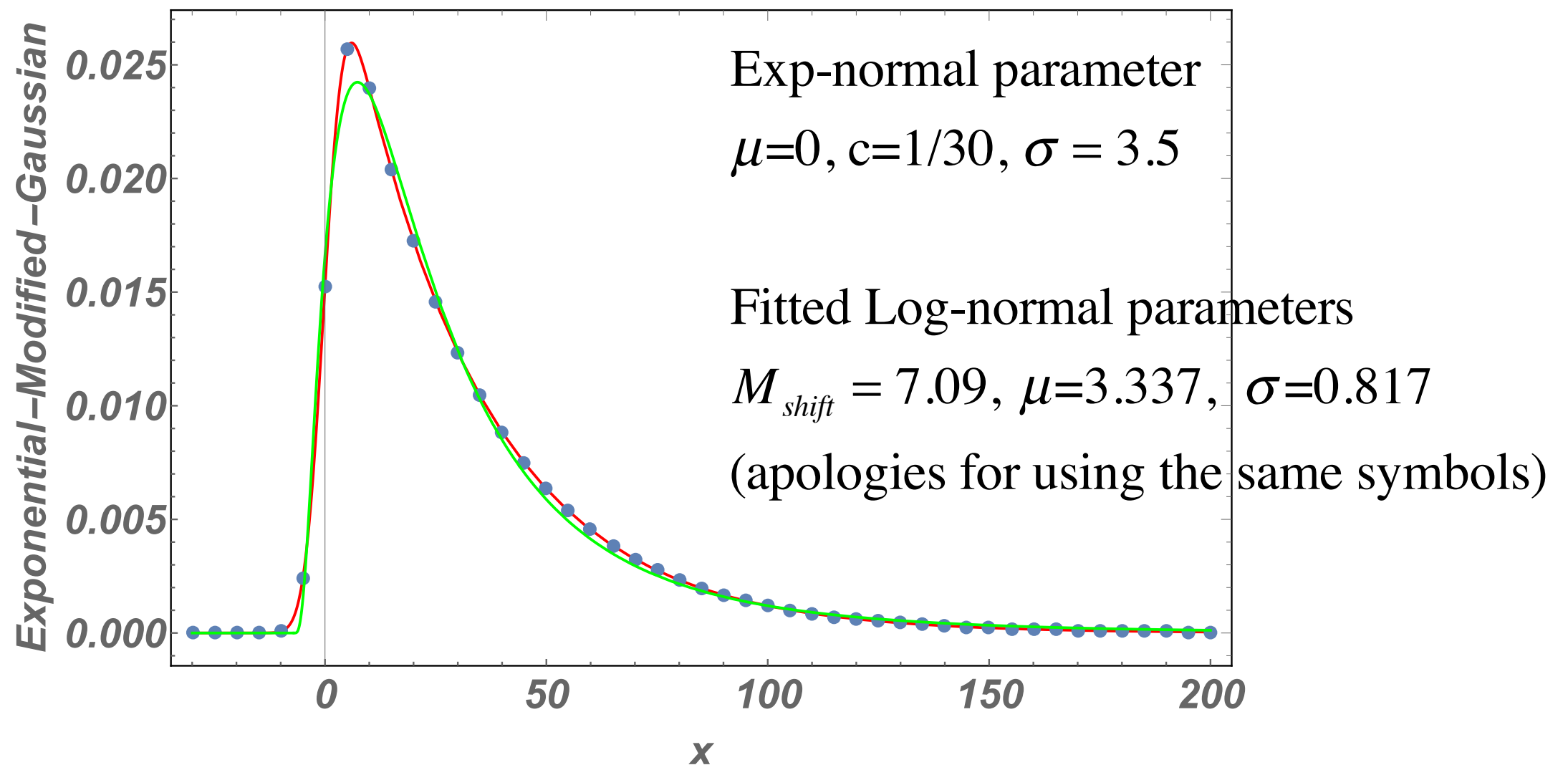


It is possible to adjust it further by a shift (M_shift)

comparison of properties

- ***Both the EMG and Lognormal have heavy tails, and therefore appear to be interesting models for data that has one sided tails.***
- ***Neither EMG and Lognormal are part of the group of Stable distributions. A distribution is called stable if a linear combination of two random variables drawn from it also has the same distribution (up to location and scale parameters)***
- ***the EMG has support over the $(-\infty, +\infty)$. Lognormal has support from 0 to ∞ (for the parameter $M=0$)***
- ***It should be possible to adjust the parameters to match the second and third moments of EMG and Lognormal to make the two look similar. The mean can then be used to adjust the shift parameter.***
- ***For experimental physics using either distribution for fitting seems reasonable, but it might be useful to understand the underlying process to see which one is a better model.***

Brute force comparison by fit



***The moment tuning can be done with an explicit fit.
The red curve is an Exp-normal with parameters in
the legend.***

References

- Scott Hanley, Ph.D. Thesis Drexel University, Practical applications and properties of the Exponentially Modified Gaussian (EMG) distribution, 2011.
- Table of integral transforms, Bateman Manuscript project, CalTech, McGraw-Hill Book company, 1954.
- Tanner Kaptanoglu, Nucl.Instrum.Meth.A 889 (2018) 69-77, 1710.03334
- Limpert, Stahel, Abbt, BioScience, May 2001, Vol. 51 no. 5. pp341
- https://en.wikipedia.org/wiki/Stable_distribution
- <https://www.sciencedirect.com/topics/engineering/lognormal-distribution>
- Wikipedia has excellent discussion of these distributions, but it is useful to derive the results by oneself.